



# Working Papers

[www.cesifo.org/wp](http://www.cesifo.org/wp)

## Jump Processes in Natural Gas Markets

Charles F. Mason  
Neil Wilmot

CESIFO WORKING PAPER NO. 4604  
CATEGORY 10: ENERGY AND CLIMATE ECONOMICS  
JANUARY 2014

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Jump Processes in Natural Gas Markets

## Abstract

Natural gas is likely to become increasingly important in the future. Understanding the stochastic underpinnings of natural gas prices will be critical, both to policy analysts and to market participants. To this end, we investigate the potential presence of jumps in natural gas spot prices in the U. S. and in the U. K. We find compelling empirical evidence for the importance of jumps in both markets, though jumps appear to appear more frequently in the U. K. Some of the difference between the U.S. and U.K. jump probabilities may be due to oil prices, other factors play a role.

JEL-Code: Q410, G170.

Keywords: natural gas prices, jump diffusion, GARCH.

*Charles F. Mason\**  
*H.A. True Chair in Petroleum and Natural  
Gas Economics, Department of Economics  
& Finance / University of Wyoming  
1000 E. University Ave  
USA – Laramie, WY 82071  
bambuzlr@uwyo.edu*

*Neil Wilmot*  
*Department of Economics  
University of Minnesota Duluth  
1318 Kirby Drive  
USA – Duluth MN 55812  
nwilmot@d.umn.edu*

\*corresponding author

An earlier version of this paper was presented at the “Oil Price Forecasts and Trends” workshop at the Fondazione Eni Enrico Mattei – Milan, in May, 2013. We thank the conference participants, and particularly John Elder and Julien Chevallerier, for their useful comments. Mason thanks the Center of Economic Studies in Munich for their kind support and hospitality.

# 1 Introduction

Over the course of the next several years, a number of important events are likely to impact the market for natural gas. Between the potentially controversial and widespread adoption of hydraulic fracturing on the one hand, and the likely move away from coal and into natural gas as an important fuel for electricity generation on the other hand, significant changes are likely to come to the fore. To accommodate these potential changes, substantial changes in the underlying infrastructure will also have to be undertaken. Power plants will have to be retrofitted or replaced, and pipeline networks will have to be expanded; both reflect large-scale investments. These investments must be contemplated in a world of profound uncertainty, in particular as regards the price path for natural gas. As such, there is a real premium of developing a better understanding of the fundamental stochastic process driving natural gas prices.

Many analysts have suggested that these prices follow a geometric Brownian motion (GBM), *i.e.*, that log returns – the inter-temporal differences in the natural log of prices – are Normally distributed (possibly allowing for some form of mean-reversion). Increasingly, however, it has been recognized that the arrival of new information can lead to unexpectedly rapid changes – or jumps – in spot prices.<sup>1</sup> The implication is that the presumption of Normally distributed log-returns may be suspect. In particular, the prospect for abnormally fat tails becomes important.

It is interesting in this regard to draw a contrast to crude oil prices. A host of papers have investigated the time series properties of oil prices, pointing to important structural

---

<sup>1</sup> Early examples of papers that argue for the use of GBM to describe oil or gas prices include Baker et al. (1998); Sadorsky (1999) and Schwartz and Smith (2000). Other authors have suggested expanding this paradigm to allow for mean-reversion; examples include Pindyck (1999) and Abadie and Chamorro (2009). Shafiee and Topal (2010) review price modeling techniques including both GBM and mean reversion, while utilizing a binary method of introducing discontinuities (jump and dips) into the process; they conclude that empirical models should include a diffusion component, mean reversion and a jump diffusion component. On the other hand, Geman (2005) has suggested that mean reversion may no longer apply, and that GBM satisfactorily represents natural gas prices, should one choose to exclude the possibility of jumps. More recent work that focuses explicitly on jumps includes Askari and Krichene (2008); Postali and Picchetti (2006); Lee et al. (2010) and Wilmot and Mason (2013). Chevallier and Sévi (2013) allow for jumps in the spot price of carbon permits in the European Union trading market.

breaks, time-varying volatility and the potential for abrupt and unexpected changes (also known as ‘jumps’).<sup>2</sup> Most likely, the attention allocated to oil prices is a reflection of the singular importance of the resource in the modern world economy. But as we noted above, natural gas seems poised to emerge as a similarly important resource, and – somewhat surprisingly – little work has been done to identify key properties of natural gas prices.

This oversight is perhaps all the more surprising when one considers the relative magnitude of variation in the two price series. In Figure 1, we plot the log returns for the three major energy resources oil, natural gas and coal – for the period from January 2004 to January 2010.<sup>3</sup>

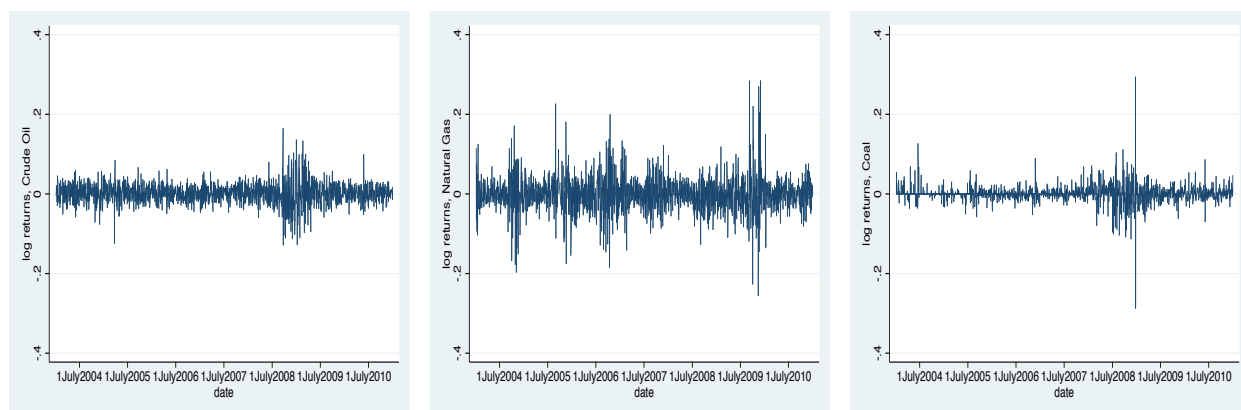


Figure 1: log price returns for 3 key energy resources: Crude oil (left panel), Natural Gas (center panel), Coal (right panel)

Two factors jump out from this comparison. First, natural gas prices are considerably more volatile than the other fuels, and second, large sudden changes in the spot price of natural gas are more frequent and larger than for the two other fuels. These observations suggest the appropriate model of the underlying stochastic process governing natural gas prices is likely to be considerably more complicated than a simple Brownian motion (BM), or geometric Brownian motion (GBM), process. The significance of this

<sup>2</sup> Most of the papers we referenced in footnote 1 consider both oil and gas, but focus mainly on oil; only Abadie and Chamorro (2009) focuses exclusively on natural gas. By contrast, several of the papers we referenced exclusively on oil.

<sup>3</sup> All plots reflect daily observations on spot prices. The specific prices used are West Texas Intermediate for oil, Henry Hub for natural gas, and Central Appalachian for coal. Date were taken from the Energy Information Administration website.

point, in turn, is underscored by the attention these two processes have received in the investment under uncertainty literature. Since the large-scale investments one imagines will be undertaken to capitalize on the apparent abundance of natural gas, and its apparent emerging demand, developing a deeper understanding of the stochastic processes driving natural gas prices would seem to have considerable importance, from both a private and a social perspective.

Our goal in this paper is to provide such an understanding. To this end, we first describe an extension of the familiar model of a stochastic process that allows for unexpected changes, or jumps. This extension leads naturally to an econometric specification, which can be readily combined with time-varying volatility (also known as the generalized autoregressive conditional heteroscedasticity, or GARCH, framework). After incorporating these elements, we characterize the likelihood function that governs the data generating process; this, in turn, leads directly to an estimation procedure and hypotheses tests regarding the appropriate specification of the stochastic process.

We then apply this econometric methodology to two important times series for natural gas prices: the spot price of natural gas at the major trading hub for natural gas in the U. S., the Henry Hub, and the spot price for natural gas in the U. K. We discuss results based on daily observations, for both spot prices. Based on these results, we use likelihood ratio tests to compare four stochastic data-generating processes: GBM, GBM allowing for a jump diffusion process (which we refer to as JD in the pursuant discussion), GBM allowing for GARCH (which we refer to as GARCH in the pursuant discussion), and GBM allowing for both GARCH and a jump diffusion process (which we refer to as GARCH-JD in the pursuant discussion). Our findings generally point to the statistical importance of allowing for both GARCH and jumps.

As both of these ingredients will induce fat tails, our empirical results may have important implications for motives to undertake large-scale investments such as major pipelines. In particular, incorporating GARCH and jumps into a model of natural gas

prices tends to increase the option value from waiting, and hence reduce the motive to undertake large scale investments in the near future.<sup>4</sup> Such a finding is consistent with the apparent sluggishness of investment into expanded pipeline export capacity out of the Marcellus shale play, for example, or the reticence of electric utilities to replace old coal-burning plants with newer gas burning facilities in the near future.

We also find a greater prevalence of jumps in European natural gas markets than in U.S. natural gas markets; indeed, the estimated probability a jump will occur in a given day is roughly an order of magnitude larger in the U. K. That differences might appear between the two natural gas price series is perhaps not surprising, given that European markets are structured quite differently to U. S. markets. In particular, there is a much greater reliance on indexing natural gas prices to crude oil prices than in the U. S., and it is conceivable that this linkage exerts an effect on spot prices. This observation naturally leads to an comparison of spot crude prices. Following this line of thought, we provide a comparison of Brent and West Texas Intermediate (WTI) spot prices. Our results suggest that while jumps are more important in European crude prices than in American crude prices, the difference in jump probabilities is substantially smaller than is the difference in jump probabilities for natural gas prices. This observation suggests other factors differentiate the two markets. We offer a brief discussion these points in the concluding remarks.

---

<sup>4</sup> Dixit and Pindyck (1993, p. 153) note that “investment is highly sensitive to volatility in project values, irrespective of investors’ or managers’ risk preferences, and irrespective of the extent to which the riskiness of  $V$  is correlated with the market”. They also observe that allowing for jumps complicates the analysis in multiple ways, by changing the expected rate of change in the underlying stochastic variate, and by increasing the percentage change in the variate over finite intervals. Wilmot (2010) uses numerical simulations to show that allowing for both positive and negative jumps tends to delay the optimal investment time, which corresponds to an increase in the option value associated with waiting to invest (Dixit and Pindyck, 1993).

## 2 Econometric Framework

In order to develop the maximum likelihood framework used to estimate the parameters of the different models, we begin with a brief examination of the stochastic processes under investigation. Let  $P_t$  denote price at time  $t$ ; its time path is said to follow a geometric Brownian motion (*GBM*) process with trend  $\alpha$  and variance parameter  $\sigma$  if

$$dP_t = \alpha P_t dt + \sigma P_t dz. \quad (1)$$

In equation (1),  $dz$  represent an increment of a Wiener process  $dz = \xi_t \sqrt{dt}$ , where  $\xi_t$  has zero mean and a standard deviation equal to 1 (Dixit and Pindyck, 1993). Denote the log returns, *i.e.*, the natural logarithm of the ratio of price in period  $t$  to the price in period  $t-1$ , by  $x_t \equiv \ln(P_t/P_{t-1})$ . If  $P_t$  follows a *GBM* process then  $x_t$  is normally distributed with variance  $\sigma^2$  and mean  $\mu \equiv \alpha - \sigma^2/2$ . This gives the pure diffusion (*PD*) model

$$x_t = \mu + \sigma z_t. \quad (2)$$

The term  $z_t$  in equation (2) is an identically and independently distributed (*i.i.d.*) random variable with mean zero and variance one.

We introduce jumps into the model in the style of Merton (1976), by assuming that two types of changes affect the log returns. The first type are ‘normal’ fluctuations, represented through the geometric Brownian motion process. The second type, ‘abnormal’ shocks, are modeled through a discontinuous process. These abnormal shocks can be thought of as occurring via the arrival of new information (Elder et al., 2013). We assume the discontinuities are described by a Poisson distribution governing the number of discrete-valued events,  $n_t \in \{0, 1, 2, \dots\}$ , that occur over the interval  $(t-1, t)$ ; accordingly,

the probability that  $j$  jumps are observed during this interval equals

$$P(N_t = j) = \frac{\exp(-\lambda)\lambda^j}{j!}. \quad (3)$$

A key element in equation (3) is  $\lambda$ , which can be interpreted as the probability of observing a jump in any brief time interval of length  $dt$ . Thus, the arrival of jumps is a Poisson distribution,<sup>5</sup> from which we can describe the change in the number of jumps observed by

$$dn_t = \begin{cases} 0 & \text{with probability } 1 - \lambda dt \\ 1 & \text{with probability } \lambda dt \end{cases} \quad (4)$$

As in Askari and Krichene (2008), when abnormal information arrives at time  $t$ , prices jump from  $P_{t-}$  (limit from left) to  $P_t = \exp(J_t)P_{t-}$ ; accordingly,  $J_t$  measures the percentage change in price. The resultant stochastic process for the random variable  $P_t$  may then be written as

$$\frac{dP_t}{P_t} = \alpha dt + \sigma dz_t + (\exp(J_t) - 1)dn_t, \quad (5)$$

where  $dz_t$  has the same properties assumed in equation (1) and  $dn_t$  is the independent Poisson process described in equation (4). Together the terms  $dz_t$  and  $dn_t$  make up the instantaneous component of the unanticipated return. It is natural to assume these terms are independent, since the first component reflects ordinary movements in price while the second component reflects unusual changes in price.

The size of the jump,  $Y_{t,k}$ , is itself a random variable; we assume it is assumed to normally distributed with mean  $\theta$  and variance  $\delta^2$ , and that it is independent of the

---

<sup>5</sup> One could of course use alternative specifications of the jump process, including Bernoulli or Levy. Our choice is motivated by the ability to combine the Poisson process – along with a GARCH process – into the basic PD econometric model. One advantage of our approach is that it leads to a relatively straightforward extension of the analytics associated with evaluating optimal investment; for example, Dixit and Pindyck (1993, p. 171) show that including a Poisson process into a conventional Brownian motion framework adds only one (non-linear) term to the key equation that determines the optimal value function associated with investing. Note too that we do not specify jump events *ex ante*, but rather let the econometric results pick out the key parameters. An alternative would be to use some criterion to decide when a jump has occurred, as in Chevallier and Sévi (2013).



distribution for the arrival of a jump. The jump component affecting returns between time  $t$  and time  $t+1$  is then

$$J_t = \sum_{k=0}^{n_t} Y_{t,k}. \quad (6)$$

Thus, the mixed jump-diffusion ( $JD$ ) process for the log-price returns can be described by

$$x_t = \mu + \sigma z_t + J_t. \quad (7)$$

An alternative explanation for the “fat tails” that are often observed in commodity price data is that  $P_t$  is subject to time-varying volatility. An example of such a phenomenon is the “generalized autoregressive conditional heteroskedastic” (GARCH) framework. Adapting the pure diffusion model to allow for this form of time-varying volatility gives the GARCH – diffusion ( $GPD$ ) process:<sup>6</sup>

$$x_t = \mu + \sqrt{h_t} z_t, \quad (8)$$

where the conditional variance,  $h_t$  is described by the process

$$h_t \equiv E_{t-1}(\sigma^2) = \kappa + \alpha_1 (x_{t-1} - \mu)^2 + \beta_1 h_{t-1}. \quad (9)$$

Note that when  $h_t = \sigma^2$  the GARCH diffusion model reduces to pure diffusion model. On the other hand, when  $\kappa > 0$  and  $\alpha_1 + \beta_1 < 1$ , the unconditional variance of the volatility of the process exists and equals  $\frac{\kappa}{1 - \alpha_1 - \beta_1}$ .

Allowing for jump discontinuities would result in the GARCH(1,1) jump-diffusion ( $GJD$ ) process:

$$x_t = \mu + \sqrt{h_t} z_t + J_t, \quad (10)$$

---

<sup>6</sup> The process described in equations (8)–(9) is characterized by four parameters,  $\mu, \kappa, \alpha_1$  and  $\beta_1$ . There is a general consensus in the literature is that a GARCH model with a limited number of terms performs reasonably well, and so we restrict our focus to this more parsimonious representation.

where  $h_t$  is described by equation (9). Duan (1997) shows that the diffusion limit of a large class of GARCH(1,1) models contain many diffusion processes allowing the approximation of stochastic volatility models by the GARCH process.

We evaluate the four models using maximum likelihood estimation methods.<sup>7</sup> To this end, we note that the parameters of our four candidate models – *PD*, *JD*, *GPD*, *GJD* – may be nested into the general log-likelihood function<sup>8</sup>

$$L(\phi, x_t) = -T\lambda - \frac{T}{2} \ln(2\pi) + \sum_{t=1}^T \ln \left[ \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \frac{1}{\sqrt{h_t + n\delta^2}} \exp\left(\frac{-(x_t - \mu - n\theta)}{2(h_t + n\delta^2)}\right) \right], \quad (11)$$

combined with the description of  $h_t$  given in equation (9). In this framework, the *GPD* model corresponds to the parameter restriction  $\lambda = \theta = \delta = 0$ ; the *JD* model corresponds to the restriction  $\alpha_1 = \beta_1 = 0$ ; and the *JD* model corresponds to the restriction  $\alpha_1 = \beta_1 = \lambda = \theta = \delta = 0$ . Comparing any pair of potential models can thus be framed as a test of an appropriate parameter restriction. For example, the comparison of the *PD* and *GPD* models is conducted by testing the parameter restriction  $\alpha_1 = \beta_1 = 0$ ; the comparison of the *PD* and *JD* models is conducted by testing the parameter restriction  $\lambda = \theta = \delta = 0$ . The empirical validity of the parameter restriction of interest can be evaluated by use of the likelihood ratio test (Johnston and DiNardo, 1997). This approach compares the likelihood function under a particular restriction,  $L(\phi^R; x)$ , to that of the unrestricted or less restricted likelihood function,  $L(\hat{\phi}; x)$ . Under the null hypothesis that the restriction is empirically valid, the decrease in the likelihood function associated with the restriction will be small. Such an approach can be used to make pairwise-comparisons between a more general model and a more restricted model. The test statistic is the log-likelihood ratio

$$LR = 2[L(\hat{\phi}; x) - L(\phi^R; x)];$$

---

<sup>7</sup> Maximum likelihood estimates are known to be consistent and invariant with asymptotically normal distributions of the parameters.

<sup>8</sup> In the empirical results we report below, the number of jumps was truncated at 10 (Ball and Torous, 1985).

under the null hypothesis this statistic will be distributed as a Chi-square random variable with  $m$  degrees of freedom, where  $m$  is the number of parameter restrictions.

### 3 Data and Its Properties

The data for this study consists of the daily closing spot price of natural gas traded at the Henry Hub in the U.S., National Balancing Point (*NBP*) in the U.K., West Texas Intermediate (*WTI*) crude oil and Brent crude oil. The oil price data and the Henry Hub spot price data is available from the Energy Information Administration’s website (U.S. Energy Information Administration, 2013a,b). U.S. spot price data are generally available over longer time frames: the *NBP* Natural Gas spot price data was obtained from Bloomberg (2013). U.S. natural gas spot price data are available early in 1997, but U.K. natural gas spot price data are only available from late summer of 2007. Similarly, the spot price of *WTI* data series begins in January of 1986, while the spot price of Brent oil begins in late May of 1987. Even focusing on the shortest series, our data still include a number of events which, *a priori*, would be good candidates for arrivals of “unusual” information: the dramatic run-up in energy prices (from late 2007 to mid 2008) and the global recession (late summer 2008), the “Arab Spring” (January – March, 2011) and superstorm Sandy (October, 2012).

For each of these time series, the price returns are calculated as

$$r_t = 100[\ln(P_t/P_{t-1})], \quad (12)$$

where  $P_t$  is the closing price on day  $t$ . A variety of summary statistics for the relative price returns of each of the time series are given in Table 1; in particular, we list the first four moments (mean, variance, skewness and kurtosis). Each series displays a substantial amount of variation; there is also evidence of asymmetry in the distribution, as displayed by the presence of the skewness. Each of the series also displays evidence of leptokurtosis

or “fat tails” by the large value for kurtosis.<sup>9</sup> We also report the Doornik-Hansen test (*DH*) test, which measures the departure from normality. Under the null hypothesis that a series is normally distributed, this statistic would be distributed as a Chi-squared random variable with 2 degrees of freedom; the values we find are substantially larger than the 1% critical value, and so the null hypothesis of a normally distributed random variable is strongly rejected for each of our time series.

Table 1: Summary Statistics, Natural Gas and Crude Oil Price Returns

	Natural Gas		Crude Oil	
	European	Henry Hub	WTI	Brent
Sample Range				
Start	10 Sept., 2007	07 Jan., 1997	02 Jan., 1986	20 May, 1987
End	30 Sept., 2013	07 Oct., 2013	07 Oct., 2013	07 Oct., 2013
Summary Statistics				
Mean	0.028	-0.0013	.020	0.026
Median	0.000	0.000	0.073	0.038
Min	-25.49	-56.82	-40.64	-36.12
Max	24.12	57.67	19.15	18.13
Std. Deviation	4.443	2.370	2.545	2.304
Variance	21.87	19.74	6.477	5.308
Coeff. of Variation	166.4	-3466	127.8	86.97
Skewness	-0.066	0.493	-0.762	-0.666
Kurtosis	8.790	23.61	17.86	17.96
DH test statistic <sup>a</sup>	782.0	9363	9380	9487
number of obs'ns	1542	4193	7004	6692

a: 1% critical value = 9.21

A histogram of realized log price changes is shown in the left panel of Figure 2 for the spot price of natural gas at the Henry Hub. For purposes of comparison, we have super-imposed a normal probability distribution with the implied mean and variance for the log changes in the spot price. If the relative prices follow a geometric Brownian motion process, the histogram should fit the probability distribution well. Noticeably, the figure points to the existence of a significant number of large changes, especially apparent in

<sup>9</sup> For a normally distributed random variable, the skewness would be 0 and the kurtosis would be 3; the statistics reported in Table 1 are broadly different from these values.

the extremes of both tails, which has the potential to influence the trigger value at which the firm holding a real option should optimally invest. The perceived leptokurtosis reinforces the inadequacy of the continuous process to fit the data, reaffirming the notion of a discontinuous process in the data-generating process. These features are corroborated by the right panel in Figure 2, which shows a “quantile-quantile” plot. This plot compares the quantiles of the empirical distribution (measured on the  $y$ -axis) against the quantiles of a theoretical normal distribution (measured on the  $x$ -axis). If the empirical distribution is close to a normal distribution, the relation will be well described by a straight line. Alternatively, if – as in this case – there are significant departures from a linear relation, then the data is not well-described by a normal distribution. In particular, departures from a linear relation in the tails indicate significant leptokurtosis.

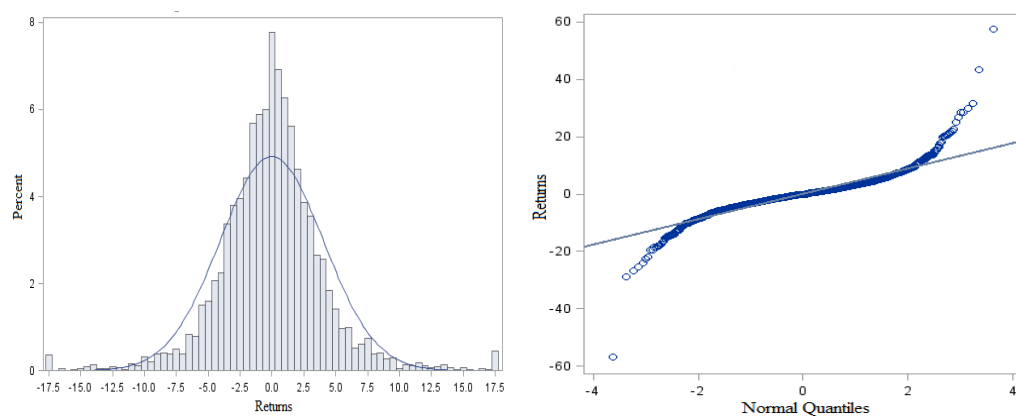


Figure 2: Histogram and fitted normal distribution (left panel) and quantile-quantile-plot (right panel) for log price returns of the Henry Hub Natural Gas

Figure 3 provides similar evidence for NBK natural gas spot prices in the U. K. As in the preceding figure we have super-imposed a normal probability distribution with the implied mean and variance for the log changes in the spot price. Here too, the figure indicates the empirical importance of events in both tails exceeds that of a Normal distribution, again reinforcing the inadequacy of the continuous process to fit the data. As

above, these features are corroborated by a “quantile-quantile” plot, here presented in the right panel in Figure 3. As above, there are significant departures from a linear relation, indicating the data is not well-described by a normal distribution, and suggesting significant leptokurtosis.

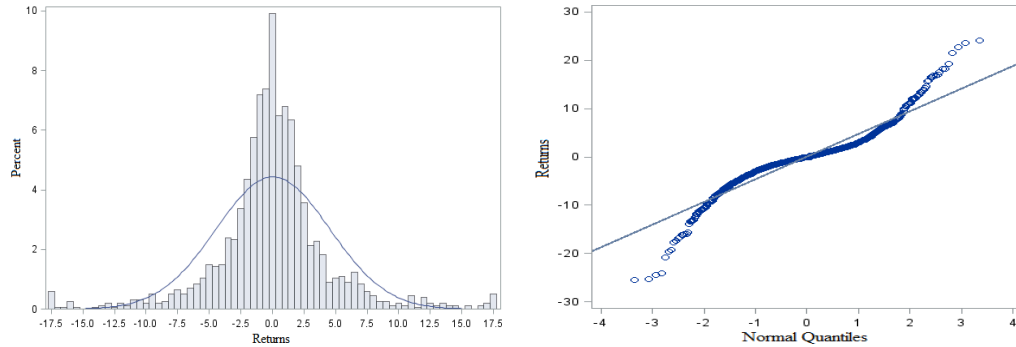


Figure 3: Histogram and fitted normal distribution (left panel) and quantile-quantile-plot (right panel) for log price returns of the UK Natural Gas

## 4 Empirical Results

### 4.1 Maximum Likelihood Estimation

Using maximum likelihood methods, we estimate the parameters for the four stochastic processes ( $PD$ ,  $JD$ ,  $GPD$ ,  $GJD$ ) using the daily observations.<sup>10</sup> We organize the presentation of these results into three tables: results for the two natural gas time series are reported in Table 2, while results for the two crude oil time series are reported in Table 3. For each of the four models and each of the price return variates, the log-likelihood statistic is given in square brackets immediately below the model acronym. For each parameter, we show the standard error in parentheses below the corresponding estimate. Statistical significance of these estimates is indicated by the use of asterisks: “\*” represents statistical significance

<sup>10</sup> Estimation was undertaken in GAUSS, utilizing the constrained maximum likelihood module. The code to estimate the GARCH model was obtained from Schoenberg (2006).

at the 10% level, “\*\*” represents statistical significance at the 5% level and “\*\*\*” represents statistical significance at the 1% level.<sup>11</sup> Test statistics for the pairwise comparisons of models are presented in Table 4, for all four series.

Table 2: Estimation of the model parameters for daily spot price returns: natural gas

model	Parameter							
	$\mu$	$\sigma$	$\kappa$	$\alpha_1$	$\beta_1$	$\lambda$	$\theta$	$\delta$
<u>European natural gas</u>								
<i>PD</i>	0.0283	4.675***						
[−2.961]	(0.038)	(0.084)						
<i>JD</i>	−0.0051	1.783***				0.5009***	0.0661	5.990***
[−2.775]	(0.033)	(0.105)			(0.059)		(0.101)	(0.377)
<i>GPD</i>	0.0775		0.4122***	0.2106***	0.7794***			
[−2.702]	(0.064)		(0.079)	(0.020)	(0.020)			
<i>GJD</i>	0.0438		0.1996**	0.2590***	0.6840***	0.1211***	−0.747	5.731***
[−2.642]	(0.056)		(0.068)	(0.003)	(0.326)	(0.032)	(0.634)	(0.764)
<u>Henry Hub</u>								
<i>PD</i>	−0.0018	4.4448***						
[−2.911]	(0.007)	(0.049)						
<i>JD</i>	−0.0680	2.8673***				0.1445***	0.4533	8.520***
[−2.756]	(0.035)	(0.069)				(0.021)	(0.488)	(0.586)
<i>GPD</i>	−0.0118		0.2843***	0.1208***	0.8692***			
[−2.711]	(0.048)		(0.038)	(0.009)	(0.009)			
<i>GJD</i>	−0.0118		0.4633**	0.1216***	0.8331***	0.0149**	2.381	12.82**
[−2.680]	(0.049)		(0.079)	(0.011)	(0.014)	(0.007)	(2.196)	(3.001)

\*: statistically different from 0 at 10% critical level

\*\* : statistically different from 0 at 5% critical level

\*\*\*: statistically different from 0 at 1% critical level

For both natural gas time series, the estimated drift parameter,  $\mu$  does not differ significantly from zero, while the estimated instantaneous rate of variance,  $\sigma$ , is statistically significant. Comparing the results for the *PD* and *JD* models, we see that allowing for

<sup>11</sup> Note that the parameters  $\sigma, \kappa, \lambda$  and  $\delta$  cannot take on negative values, and so the hypothesis tests for these parameters are one-tailed. For the other four parameters, the hypothesis tests are two-tailed.

jumps tends to markedly reduce the estimated instantaneous rate of variance. We also note that allowing for jumps improves the model's predictive power, as indicated by the increase in the log-likelihood function. Comparing the results for the *PD* and *GPD* models, we see allowing for GARCH also improves the model's predictive power – here too the log-likelihood function is increased. In addition, all parameters associated with the GARCH process are strongly statistically significant. Contrasting *JD* with *GJD*, we see that including GARCH in a model with jumps reduces the estimated mean value of the jump, which also becomes statistically indistinguishable from zero. But at the same time, the probability of a jump occurring, and the estimated variance of the jump size, remain statistically significant. Thus, while allowing for GARCH evidently captures some of the estimated effect of the jumps in the *JD* model, it does not eliminate their role. Finally, comparing the *GPD* and *GJD* models, we see that allowing for both GARCH and jumps generates further improvements in predictive power. Inspecting the test statistics for the various model comparison, we see that the improvement in fit associated with adding jumps or GARCH to the *PD* model is statistically important, as the corresponding test statistics are substantially larger than the 1% cutoff values. Moreover, combining GARCH with jumps yields a statistically significant improvement over the *JD* model and the *GPD* model. We conclude that the statistically preferred model has both GARCH and jumps.

Turning next to crude oil price returns, we see that – similar to the natural gas price returns – allowing for jumps reduces the estimated variance. We also note the broad significance of all parameters in the GARCH model, as well as the significance of estimated jump probability and variance of jump sizes; these results mirror those for natural gas. Also parallel to the natural gas results, allowing for jumps or GARCH generates statistically important increases in predictive power over the *PD* model. Also parallel to the natural gas results, combining GARCH with jumps yields a statistically significant improvement over the *JD* model; similarly, combining jumps with GARCH yields a statistically significant improvement over the *GPD* model. Again, we conclude



that the statistically preferred model has both GARCH and jumps.,

Table 3: Estimation of the model parameters for daily spot price returns: crude oil

model	Parameter							
	$\mu$	$\sigma$	$\kappa$	$\alpha_1$	$\beta_1$	$\lambda$	$\theta$	$\delta$
<u>West Texas Intermediate</u>								
<i>PD</i>	-0.9385	2.552***						
[−2.353]	(0.035)	(0.022)						
<i>JD</i>	-0.0872***	1.712***			0.1669***	-0.404***	4.469***	
[−2.239]	(0.026)	(0.035)			(0.021)	(0.170)	(0.247)	
<i>GPD</i>	0.0298		0.0802***	0.0946***	0.8954***			
[−2.199]	(0.023)		(0.009)	(0.008)	(0.008)			
<i>GJD</i>	0.0610***		0.0493***	0.0527***	0.9229***	0.0464***	-0.905	5.129***
[−2.172]	(0.023)		(0.009)	(0.006)	(0.007)	(0.009)	(0.386)	(0.499)
<u>Brent</u>								
<i>PD</i>	0.0265	2.304***						
[−2.253]	(0.015)	(0.019)						
<i>JD</i>	0.0562	1.650***			0.1568***	-0.190	3.902***	
[−2.162]	(0.035)	(0.069)			(0.021)	(0.488)	(0.586)	
<i>GPD</i>	0.0494**		0.0609***	0.0804***	0.9096***			
[−2.123]	(0.022)		(0.007)	(0.007)	(0.007)			
<i>GJD</i>	0.0695***		0.0297***	0.0540***	0.9252***	0.0779***	0.415*	3.530***
[−2.106]	(0.023)		(0.008)	(0.006)	(0.007)	(0.021)	(0.236)	(0.443)

\*: statistically different from 0 at 10% critical level

\*\* : statistically different from 0 at 5% critical level

\*\*\*: statistically different from 0 at 1% critical level

A comparison of the results for European and American natural gas spot price returns is particularly interesting. As *GJD* is the statistically preferred model, we focus on these results. First, we note that the estimated value of  $\kappa$  is larger for the Henry Hub data, while the estimated values of  $\alpha_1$  and  $\beta_1$  are larger for the UK data. This would seem to suggest a greater degree of persistence in time-varying volatility in European markets. At the same time, the estimated probability of a jump occurring is roughly an order of magnitude larger in the UK than at the Henry Hub. This important difference might point

to the importance of relatively thicker markets in the US than in the UK, which could serve to buffer some shocks. This distinction between the two markets could also explain the relatively larger persistence of time-varying volatility in the European markets.

Table 4: Likelihood ratio test comparisons of the models<sup>a,b</sup>

time series	$LR_{PD,JD}$	$LR_{PD,GPD}$	$LR_{JD,GJD}$	$LR_{GPD,GJD}$
<u>Natural Gas</u>				
European spot returns	573.5	799.4	409.9	184.0
Henry Hub spot returns	1296	1670	635.2	260.3
<u>Crude Oil</u>				
WTI spot returns	1595	2157	939.57	377.5
Brent spot returns	1229	1743	742.7	229.4

*a*: 1% critical value for  $LR_{PD,JD}$ ,  $LR_{GPD,GJD}$  = 11.34

*b*: 1% critical value for  $LR_{PD,GPD}$ ,  $LR_{JD,GJD}$  = 9.21

An important institutional feature of European markets is their tendency to index long-term natural gas prices to crude oil prices; this indexing is not generally used in the US. This observation raises the question: might the reliance on crude oil price indexing in long-term contracting spill over into spot prices? If so, could the increased jump probability and larger persistence of time-varying volatility in the UK natural gas returns simply be a reflection of greater tendency for jumps, and more persistence in time-varying volatility, in European crude oil markets vis-a-vis American crude oil markets. Comparing the estimated jump probabilities in WTI and Brent price returns, we see that there is in fact a larger tendency for jumps in the European market. But this increase is more like a doubling for crude oil markets, as opposed to the order of magnitude increase for natural gas. Apparently there are other forces at play, perhaps related to market thickness – as we argued above.

## 4.2 Robustness Checks

The results we presented above were obtained using the longest available time series available at the time, variate by variate. For crude oil spot prices, these data go back in the 1980s: January, 1986 for U.S. prices, May, 1987 for European prices. By contrast, natural gas spot price data are available for a much shorter time frame: since January, 1997 for U.S. prices, and only since September, 2007 for European spot prices. The heterogeneity in the length of these time series does not pose any problems in terms of identifying the appropriate econometric model, nor in terms of interpreting the role played by jumps and GARCH. But it does complicate the comparisons across energy price time series. In particular, since the time series for European natural gas spot prices are so much shorter than for U.S. spot prices, it is fair to ask if the significant difference between estimated jump probabilities is the result of different time periods.<sup>12</sup> To address this concern, we re-estimated our models using the largest available common sample. Since the shortest sample used in the results above is associated with the U. K. natural gas spot prices, that time frame (10 September, 2007 to 30 September, 2010) defines the largest available common sample. We therefore use this time frame for our common sample. Estimation results based on data from this common sample are reported in Tables 5 and 6.

Because the sample we used here corresponds to the sample used to estimate the various models for the U. K. natural gas data, the results reported for that variate in Table 5 are identical to the results reported for that variate in Table 2. But with the somewhat different sample utilized here, we see that the results for the U. S. natural gas variate, *i.e.*, Henry Hub prices, are a bit different to the results reported above. That said, we note that these results are qualitatively similar to the results reported earlier. In particular, the point estimates are broadly similar across the two samples. This aspect of the analysis is comforting, as it supports the implicit assumption that the time series did not experience

---

<sup>12</sup> For that matter, one may wonder if the pressures in play were substantially different in the late 1990s as opposed to the period after 2007.

a structural break at some point during the longer time frame. And referring to the test statistics for the various pairwise-comparisons of the models, reported in Table 7, we see that allowing for jumps or GARCH renders a statistically important improvement over the simple PD model, as before. Further, combining GARCH and jumps yields a statistically important improvement over both the JD and GPD models – *i.e.*, the combined model is better able to explain the data, an effect that is statistically persuasive with very strong significance. However, while the estimated jump probability in the preferred model (the one combining GARCH and jumps) is of a similar magnitude for the two samples, it is not statistically significant for the shorter sample. This may simply reflect there is less data, but a more likely explanation is that the shorter period includes some singularly unique observations; with fewer observations in the sample, these observations inject relatively more noise, which would explain the significantly larger standard error associated with the estimated jump probability. The unique observation occurred on 12 September, 2009, when a key pipeline servicing the southern U. S. issued an “Operational Flow Order” (OFO) that dramatically and adversely impacted the spot market, driving prices down from there already low level.<sup>13</sup> Looking past the statistical importance of the estimate jump probability, then, we believe there is still reason to believe that European natural gas markets are characterized by a more substantial jump probability; in turn, this implies

---

<sup>13</sup> We are most grateful to Gavin Roberts for identifying this event, and locating the following quote issued by the carrier in question, Southern Natural Gas:

“Based on historically high storage inventory levels and projected high injection requirements for the upcoming weekend, this is to inform you that Southern is implementing an OFO Type 6 for long imbalances on Southern’s contiguous pipeline system effective for the start of the gas day, Saturday, September 12, 2009.”

The particular order Southern issued imposed stark penalties for under-delivery of gas, relative to previously scheduled levels. This had the immediate effect of convincing (some, perhaps many) traders that there was a distinct possibility of an imminent collapse in natural gas markets; this led to a dramatic drop in prices. When it became clear the next day that traders had over-reacted, prices adjusted upwards, again very dramatically. Accordingly, we observed two very large log price returns on these two days, one in each direction; these returns are markedly larger than the log returns observed for the remainder of the full sample on Henry Hub spot prices. We do not make this point to argue the observations should be dropped, but rather to explain how they could play a more pronounced role in the shorter sample, in particular sharply increasing the standard error associated with the estimated jump probability. With this significant increase in the standard error, but essentially no change in the point estimate, it is perhaps not surprising that the estimate was rendered statistically insignificant.

more frequent arrivals of jumps.

Table 5: Estimation of the model parameters for daily spot price returns, common sample: natural gas

model	Parameter							
	$\mu$	$\sigma$	$\kappa$	$\alpha_1$	$\beta_1$	$\lambda$	$\theta$	$\delta$
<u>European natural gas</u>								
<i>PD</i>	0.0283	4.675***						
[−2.961]	(0.038)	(0.084)						
<i>JD</i>	−0.0051	1.783***				0.5009***	0.0661	5.990***
[−2.775]	(0.033)	(0.105)				(0.059)	(0.101)	(0.377)
<i>GPD</i>	0.0775		0.4122***	0.2106***	0.7794***			
[−2.702]	(0.064)		(0.079)	(0.020)	(0.020)			
<i>GJD</i>	0.0438		0.1996**	0.2590***	0.6840***	0.1211***	−0.747	5.731***
[−2.642]	(0.056)		(0.068)	(0.003)	(0.326)	(0.032)	(0.634)	(0.764)
<u>Henry Hub</u>								
<i>PD</i>	−0.0304	3.679***						
[−2.722]	(0.036)	(0.066)						
<i>JD</i>	−0.1046	2.687***				0.0955***	0.7747	8.003***
[−2.609]	(0.094)	(0.108)				(0.033)	(0.889)	(1.24)
<i>GPD</i>	−0.0413		0.1235***	0.0753***	0.9147***			
[−2.554]	(0.071)		(0.029)	(0.012)	(0.012)			
<i>GJD</i>	−0.0522		0.1282**	0.0617***	0.9208***	0.0123	−0.0006	9.859***
[−2.550]	(0.071)		(0.058)	(0.011)	(0.015)	(0.013)	(0.287)	(4.99)

\*: statistically different from 0 at 10% critical level

\*\* : statistically different from 0 at 5% critical level

\*\*\*: statistically different from 0 at 1% critical level

The persistently larger jump probability for U. K. natural gas prices is noteworthy, and as above can not be fully explained by effects in crude markets. Perusing the results for crude oil spot prices with the shorter common sample, reported in Table 6, we see that the difference between estimated jump probabilities in the U. S. and European markets is more pronounced for the shorter sample. While the ratio between the estimated probability for Brent spot prices and the estimated probability for WTI spot prices is on the order of 6,

whereas it was more like 2 for the longer samples, this ratio is still smaller than the ratio of jump probabilities for natural gas markets. The sense one takes from these numbers is that while differences in the movements of crude spot prices between Europe and the U. S. do seem to play some role in the heterogeneities between European and U. S. natural gas spot prices, there is more going on.

Table 6: Estimation of the model parameters for daily spot price returns, common sample: crude oil

model	Parameter							
	$\mu$	$\sigma$	$\kappa$	$\alpha_1$	$\beta_1$	$\lambda$	$\theta$	$\delta$
<u>West Texas Intermediate</u>								
<i>PD</i>	0.018***	2.569***						
[−2.363]	(0.035)	(0.022)						
<i>JD</i>	0.081	1.668***				0.1833***	−0.3426	4.558***
[−2.249]	(0.057)	(0.064)				(0.037)	(0.316)	(0.432)
<i>GPD</i>	0.0754		0.0677***	0.0852***	0.9038***			
[−2.161]	(0.046)		(0.028)	(0.017)	(0.019)			
<i>GJD</i>	0.0447		0.0325**	0.0535***	0.9331***	0.0196*	0.6287	5.207***
[−2.145]	(0.047)		(0.016)	(0.012)	(0.015)	(0.014)	(1.702)	(1.587)
<u>Brent</u>								
<i>PD</i>	0.024	2.262***						
[−2.235]	(0.036)	(0.041)						
<i>JD</i>	0.0635	1.494***				0.2255***	−0.1755	3.4934***
[−2.140]	(0.052)	(0.097)				(0.078)	(0.195)	(0.491)
<i>GPD</i>	0.0702		0.0336***	0.0497***	0.9403***			
[−2.066]	(0.043)		(0.006)	(0.008)	(0.008)			
<i>GJD</i>	0.1287**		0.0034	0.0406***	0.9494***	0.1315**	−0.6271	2.146***
[−2.056]	(0.061)		(0.009)	(0.007)	(0.007)	(0.075)	(0.340)	(0.509)

\*: statistically different from 0 at 10% critical level

\*\* : statistically different from 0 at 5% critical level

\*\*\*: statistically different from 0 at 1% critical level

Table 7: Likelihood ratio test comparisons of the models, common sample<sup>a,b</sup>

time series	$LR_{PD,JD}$	$LR_{PD,GPD}$	$LR_{JD,GJD}$	$LR_{GPD,GJD}$
<u>Natural Gas</u>				
European Spot Returns	573.5	799.4	409.9	184.0
Henry Hub Spot Returns	344.0	510.9	179.9	13.04
<u>Crude Oil</u>				
WTI Spot Returns	346.1	615.7	318.5	48.89
Brent Spot Returns	288.9	512.5	254.6	30.99

a: 1% critical value for  $LR_{PD,JD}$ ,  $LR_{GPD,GJD}$  = 11.34

b: 1% critical value for  $LR_{PD,GPD}$ ,  $LR_{JD,GJD}$  = 9.21

## 5 Conclusion

There are many reasons why a better understanding of the stochastic process driving the prices of natural gas and oil would be useful. These energy resources can have important microeconomic effects, with commodity price risk having a potentially significant impact on profits in a variety of lines of business. Knowledge of the underlying stochastic behavior of these assets could aid in forecasting spot prices, with attendant reductions in risk exposure. Moreover, decisions to invest in important infrastructure can be improved by an enhanced understanding of the stochastic processes driving the prices of related resource. For example, the accuracy of a decision to significantly expand pipeline infrastructure, or to build an export terminal for liquified natural gas, will almost surely be improved by such enhanced understanding. This is particularly true when oil and gas prices are subject to infrequent jumps, as our results indicate. For in this case, the underlying distribution of oil prices is “fat-tailed” or leptokurtotic, and fat tails can be particularly important if prices exert a non-linear marginal impact on the agent’s profit flow (Weitzman, 2009).

Furthermore, events in the market for key energy resources such as natural gas and crude oil can be important for policymakers since price shocks in these key resources are often followed by economic downturns (Hamilton, 2013).

The point of this paper is to re-examine the assumption that the relative price returns of natural gas and crude oil can be modeled using a continuous time process, with an eye towards developing a more accurate understanding of the stochastic forces driving these spot prices. We draw several important conclusions from our analysis. For both natural gas and crude oil prices, in both the US and Europe, the data strongly suggest that allowing for jumps or time-varying volatility generates improved fit, relative to the pure diffusion model. Moreover, it is apparent that combining jumps with a GARCH process with jumps (GJD) outperforms all alternative stochastic processes. Thus, our results indicate that incorporating both time-varying volatility and jumps into empirical models of natural gas spot prices improves predictive power; the sharper predictions that result from this improvement should be of clear benefit to market traders. But we believe there are further benefits, associated with a more accurate description of the stochastic processes driving natural gas prices. These added benefits accrue from the ability to better frame the underlying stochastic model in an investment under uncertainty framework, which we believe has real potential for evaluating important large-scale infrastructure investments such as pipelines or liquified natural gas export terminals. Because such enhancements to transportation infrastructure may have far-reaching benefits, for example by facilitating gas movements to regions with larger demand, the welfare consequences of these investments may be substantial. The potential for substantial welfare implications of these investments underscores the importance of developing a better understanding of the stochastic process underlying natural gas prices, which in turn highlights the value of developing a more accurate empirical model to describe these prices.



## References

- Abadie, L. M. and Chamorro, J. M. (2009). Monte carlo valuation of natural gas investment, *Review of Financial Economics* **18**: 10 – 22.
- Askari, H. and Krichene, N. (2008). Oil price dynamics (2002 – 2006), *Energy Economics* **30**: 2134–2153.
- Baker, M., Mayfield, S. and J., P. (1998). Alternative models of uncertain commodity prices for use with modern asset pricing methods, *The Energy Journal* **19**: 115 – 148.
- Ball, C. and Torous, W. N. (1985). On jumps in common stock prices and their impact on call option pricing, *Journal of Finance* **40**: 155–173.
- Bloomberg (2013). Natural gas spot price data for National Balancing Point, September 10, 2007 – September 30, 2013, via Bloomberg LP, accessed October 10, 2013.
- Chevallier, J. and Sévi (2013). On the stochastic properties of carbon futures prices, *Environmental and Resource Economics* **in press**.
- Dixit, A. and Pindyck, R. (1993). *Investment under uncertainty*, Princeton University Press, Princeton, NJ.
- Duan, J. C. . (1997). Augmented garch( $p,q$ ) process and its diffusion limit, *Journal of Econometrics* **79**: 97–127.
- Elder, J., Miao, H. and Ramchander, S. (2013). Jumps in oil prices: The role of economic news, *The Energy Journal* **34**: 217–237.
- Geman, H. (2005). Energy commodity prices: Is mean-reversion dead?, *Journal of Alternative Investments* **8**: 31 – 45.

- Hamilton, J. (2013). Historical oil shocks, in R. E. P. Parket and R. Whaples (eds), *Routledge Handbook of Major Events in Economic History*, Routledge Taylor and Francis Group, New York, NY, pp. 239–265.
- Johnston, J. and DiNardo, J. (1997). *Econometric Methods*, McGraw – Hill Companies Inc., New York, NY.
- Lee, Y., Hu, H. and Chiou, J. (2010). Jump dynamics with structural breaks for crude oil prices, *Energy Economics* **32**: 343 – 350.
- Merton, R. (1976). Option pricing when the underlying stock returns are discontinuous, *Journal of Financial Economics* **3**: 125–144.
- Pindyck, R. S. (1999). The long-run evolution of energy prices, *The Energy Journal* **20**: 1 – 27.
- Postali, F. and Picchetti, P. (2006). Geometric brownian motion and structural breaks in oil prices: A quantitative analysis, *Energy Economics* **28**: 506 – 522.
- Sadorsky, P. (1999). Oil price shocks and stock market activity, *Energy Economics* **21**: 449 – 469.
- Schoenberg, R. (2006). Generalized autoregressive conditional heteroskedasticity models. <http://faculty.washington.edu/rons/garch.html>; accessed 08/26/2006.
- Schwartz, E. S. and Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices, *Mangement Science* **47**: 893 – 911.
- Shafiee, S. and Topal, E. (2010). A long-term view of worldwide fossil fuel prices, *Applied Energy* **87**: 988 – 1000.
- U.S. Energy Information Administration (2013a). Natural gas spot and future prices. <http://www.eia.gov/naturalgas>; accessed 12/03/2013.

U.S. Energy Information Administration (2013b). Spot prices.  
<http://www.eia.gov/petroleum>; accessed 12/03/2013.

Weitzman, M. (2009). On modeling and interpreting the economics of catastrophic climate change, *Review of Economics and Statistics* **91**: 1–19.

Wilmot, N. A. (2010). *Investment under Uncertainty Utilizing Alternative Stochastic Processes*, PhD thesis, University of Wyoming, Department of Economics & Finance.

Wilmot, N. A. and Mason, C. F. (2013). Jump processes in the market for crude oil, *The Energy Journal* **34**: 33 – 48.